Final Exam Discussion

Logistics

- Mon, Dec 12th, 7-9pm
- Take-home exam in CMS
  - Available shortly before 7PM
  - Scanned submissions accepted
  - Late submissions accepted
    - for a few hours, appropriate late penalty
- There will be office hours next week
  - Consult schedule in CMS
- See me ASAP if you have a conflict!

Exam Details

- Cumulative
- Emphasis on material not in the prelim
- Time is very limited
  - so some material will end up not being covered
  - many questions will be short-answer

Exam Details

- BUT I may still ask things like:
  - Write a query in SQL or Relational Algebra
  - Solve an optimization/cost calculation question
  - Questions on indexing
  - … and other stuff from early in the course

Recommendation

- Go over all the homeworks!!
  - And the prelim
- You should be familiar with **everything that was in any homework**, even if you chose to split up work so that your partners did a particular question…

New topics for Final (not in Prelim)

- Functional Dependencies/ER Diagrams
- Concurrency and Recovery
- Distributed Systems Topics
  - 2PC, Paxos, Eventual Consistency, …
- Column Stores
- MapReduce

ER Diagrams

- Be familiar with the ER diagrams introduced in class and in the textbook
  - Including extended features like weak entities
- Understand how to translate/map between ER diagrams and relational tables

ER diagrams

- Typical questions you can expect:
  - Draw an ER diagram for this real-world scenario
  - Convert this ER diagram to a SQL schema or vice versa

Functional dependencies

- Understand why redundancy is bad and how FDs relate to redundancy
- Understand FDs
  - Definition
  - Relationship to keys
  - Armstrong's Axioms, soundness, completeness etc
  - What a closure, attribute closure and cover are
  - Be able to do proofs on the above

Decompositions and Normal Forms

- Understand problems with decomposition
  - Performance hit
  - May be lossy
  - May not be dependency preserving
- Know definitions of BCNF and 3NF, how to check if a relation is in one of these, and how to decompose

Concurrency and Transactions

- ACID properties
  - Atomicity
  - Durability
  - Consistency
  - Isolation

Isolation/concurrency

- Formal/theoretical concepts:
  - Isolation anomalies
  - Various kinds of serializability
  - Recoverable/ACA/strict schedules

Locking Protocols

- 2PL and its variants (conservative, strict, both)
  - How they work
  - Pros and cons
- Advanced/custom locking protocols
  - Phantoms and index locking
  - Locking in B+-trees
  - Multiple granularity locks (IS, IX, S, X)
- Deadlock and solutions

Nonlocking protocols

- Optimistic
- Version-based
  - Timestamp-based multiversion CC
  - Snapshot isolation

SQL isolation levels

Level	Dirty Read	Unrepeatable Read	Phantom
READ UNCOMMITTED	Possible	Possible	Possible
READ COMMITTED	No	Possible	Possible
REPEATABLE READ	No	No	Possible
SERIALIZABLE	No	No	No

Recovery and ARIES

LogRecords

- prevLSN
- XID
- type
- pageID
- length
- offset
- before-image
- after-image

Data pages

each with a pageLSN

master record

Xact Table

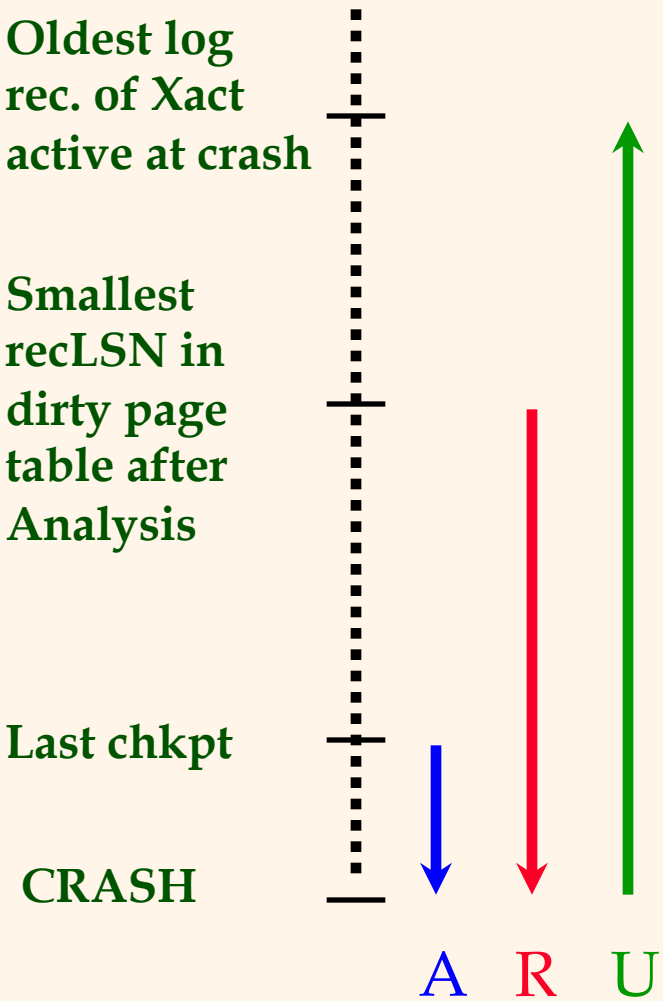
- lastLSN
- status

Dirty Page Table

- recLSN

flushedLSN

Crash Recovery: Big Picture



- Start from a checkpoint (found via master record).
- Three phases. Need to:
  - Figure out which Xacts committed since checkpoint, which failed (Analysis).
  - REDO *all* actions.
    - (repeat history)
  - UNDO effects of failed Xacts.

Distributed systems

- Query processing
  - New join algorithms (semijoin, Bloomjoin)
- Consensus protocols
  - 2PC and variants
  - 3PC
  - Paxos

Distributed systems

- Eventual Consistency
  - Understand Vector Clocks
- Spanner
  - Commit time via 2PC
- Basically at the short-answer question level

Column Stores

- Performance advantages for OLAP (vs OLTP)
- Basically at the short-answer question level

MapReduce

- Understand basic processing model
- May ask you to provide pseudocode to solve a particular problem in Map Reduce
- Look at the examples
  - especially graph processing
- Common error: solving everything in one mapper/reducer, i.e. not using parallelism

That's All Folks!

- See you in office hours …
- Have a great semester break!